

# 基于改进 ConvNeXt 和知识蒸馏的蘑菇图像识别方法<sup>\*</sup>

任喜伟 王 瑞 贾事端 刘 艳 肖 曼

(陕西科技大学电子信息与人工智能学院 西安 710021)

**摘 要:** 蘑菇种类繁多,尤其有毒蘑菇形态相近,不易识别,高效识别蘑菇种类有重要的现实需求。针对现有蘑菇图像识别方法存在背景复杂,识别精度不高,模型参数量大,移动端部署困难的问题,提出了一种基于改进 ConvNeXt 模型和知识蒸馏的蘑菇图像识别方法。首先,通过迁移学习将预训练的 ConvNeXt 权重文件应用于蘑菇识别任务,并引入坐标注意力机制,构建了 ConvNeXt\_CA 模型,有效提升模型的细粒度特征提取能力。其次,基于知识蒸馏技术,将 ConvNeXt\_CA 模型作为教师模型,ShuffleNet v2 模型作为学生模型,构建了轻量化 MushNet 模型。极大提升改进模型部署边缘端的整体效率。最后,进行了相关模型对比实验,结果表明,提出的改进模型准确率达到 94.89%,知识蒸馏后的 MushNet 模型大小约为原来的 1/21,均优于其他传统模型和轻量化模型。证明了所提蘑菇图像识别方法的有效性和可行性。

**关键词:** 蘑菇图像识别;MushNet;ConvNeXt;坐标注意力机制;迁移学习;知识蒸馏

**中图分类号:** TN911.73;TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Mushroom image recognition method based on improved ConvNeXt and knowledge distillation

Ren Xiwei Wang Rui Jia Shiduan Liu Yan Xiao Man

(School of Electronic Information and Artificial Intelligence, Shaanxi University of Science & Technology, Xi'an 710021, China)

**Abstract:** There are many kinds of mushrooms, especially poisonous mushrooms, which are similar in shape and difficult to identify. There is an important practical need for efficient identification of mushroom species. In view of the problems of complex background, low recognition accuracy, large number of model parameters and difficult deployment on mobile terminals in existing mushroom image recognition methods, a mushroom image recognition method based on improved ConvNeXt model and knowledge distillation is proposed. Firstly, the pre-trained ConvNeXt weight file is applied to the mushroom recognition task through transfer learning, and the coordinate attention mechanism is introduced to construct the ConvNeXt\_CA model, which effectively improves the fine-grained feature extraction ability of the model. Secondly, based on the knowledge distillation technology, the ConvNeXt\_CA model is used as the teacher model and the ShuffleNet v2 model is used as the student model to construct a lightweight MushNet model. The overall efficiency of the edge deployment of the improved model is greatly improved. Finally, the relevant model comparison experiments are carried out. The results show that the accuracy of the proposed improved model reaches 94.89%, and the size of the MushNet model after knowledge distillation is about 1/21 of the original, which is better than other traditional models and lightweight models. The effectiveness and feasibility of the proposed mushroom image recognition method are proved.

**Keywords:** mushroom image recognition; MushNet; ConvNeXt; coordinate attention mechanism; transfer learning; knowledge distillation

## 0 引 言

蘑菇富含蛋白质和氨基酸,是全球广泛食用的食材<sup>[1]</sup>。

我国已知蘑菇种类逾 1 400 种,其中近 70%可食用,其余多具毒性,误食可能导致严重后果<sup>[2]</sup>。因此,开发高效、准确的蘑菇识别方法对保障食品安全具有重要意义。传统蘑菇

识别依赖形态观察、化学检测等手段,存在效率低、主观性强等问题,难以适应大规模应用需求<sup>[3]</sup>。近年来,机器学习与深度学习的发展为蘑菇图像自动识别提供了新路径,尤其在处理复杂图像方面展现出显著优势,逐渐成为研究热点。

近几年已有研究表明,深度学习在蘑菇识别领域具有广阔的应用前景。Peng 等<sup>[4]</sup>在 MobilenetV2 网络中引入压缩和激励注意力机制(squeeze and excitation,SE),设计多维注意力模块,结合空洞空间卷积池化金字塔(atrous spatial pyramid pooling,ASPP)模块捕获长距离依赖关系,在 mushroom 和 MO106 数据集上验证了模型的有效性;CENGIL 等<sup>[5]</sup>使用预训练的 YOLOv5 算法,在进行参数微调的改进下,对 8 种 644 张有毒蘑菇图像进行分类识别;王莉等<sup>[6]</sup>为解决干香菇等级识别技术复杂及识别精度不高的问题,在 ResNet18 骨干网络中引入高效通道注意力机制;姚芷馨等<sup>[7]</sup>提出 YWeight 权重衰减方法,通过控制有效学习率优化模型泛化性能;张志刚等<sup>[8]</sup>提出了一种改进的 ResNeXt50 网络,设计了一种多尺度特征引导模块,采用改进的注意力机制模块来减少网络对复杂背景的学习,对模型中的不同层级特征进行融合,利用得到的联合特征进行识别。

尽管现有方法在识别准确性和自动化方面有所提升,但仍存在一些局限性:轻量级网络虽然计算效率较高,但其特征提取能力有限,在复杂背景或多尺度变化下难以获得理想的识别精度;ResNet、ResNeXt 等深层卷积网络虽然具有较强的特征提取能力,但其复杂的堆叠结构与跳跃连接增加了模型优化与解释的难度,同时在全局信息捕获与多层次特征融合方面仍显不足,难以有效应对细粒度识别任务;现有方法在模型泛化性与鲁棒性方面仍有待加强,特别是在面对样本分布不均衡或噪声干扰较大的实际场景时,容易出现过拟合或误判现象。

为此,本文选择一种兼具局部感知能力与全局建模优势的新型主干网络——ConvNeXt<sup>[9]</sup>作为基础架构。该网络借鉴视觉变换器(vision transformer, ViT)的设计理念,在保留卷积操作局部感受野优势的同时,通过改进的块结构增强了长距离依赖关系的建模能力,具有良好的分类性能与计算效率。

在此基础上,本文提出一种改进模型——ConvNeXt\_CA,在 ConvNeXt 的 block 模块中融入坐标注意力(coordinate attention,CA)<sup>[10]</sup>模块,使模型能够更有效地聚焦于图像关键区域,降低背景噪声干扰,从而提升模型表现。为了充分利用预训练模型所具备的通用表征能力,同时有效训练本文引入的 CA 模块,采用了参数冻结、模块解冻的迁移学习<sup>[11]</sup>策略。通过解冻新增模块,冻结其他网络层,使新增模块在预训练特征空间中自适应调整参数,以充分发挥其对蘑菇图像识别的细粒度特征提取能力。同时,采用知识蒸馏<sup>[12]</sup>,将计算量高、结构复杂的 ConvNeXt\_CA

作为教师网络,将其知识迁移至轻量级网络 ShuffleNetV2<sup>[13]</sup>中,得到轻量化蘑菇识别网络 MushNet,为蘑菇图像识别方法部署边缘设备端提供可能。

## 1 本文方法

### 1.1 教师模型设计

尽管 ConvNeXt 在多种视觉任务中展现了出色的性能,并具备良好的泛化能力,但其局部特征提取方面仍有不足,通道间特征融合不够充分,导致在处理细节丰富或复杂背景下的图像时可能出现特征混淆。此外,其计算开销较大也限制了其在资源受限设备上的部署与应用。

针对上述问题,已有研究从不同角度对 ConvNeXt 进行改进。杨宇轩等<sup>[14]</sup>通过减少网络层数并引入 Ghost 卷积降低模型复杂度,结合简单、无需参数的注意力模块(simple, strong and effective attention module, SimAM)提升性能;杨德龙等<sup>[15]</sup>将  $7 \times 7$  卷积分解为多个小核并引入 SE 模块,增强局部特征提取与通道融合能力;田甜等<sup>[16]</sup>则加入 SimAM 提升模型性能,并采用 Focal Loss 缓解类别不平衡问题。

尽管上述方法提升了模型性能,但是针对蘑菇识别任务仍存在如下问题:Ghost 卷积难于抑制背景干扰;小核卷积感受野受限,影响整体特征建模;SimAM 的空间建模能力较弱。此外,多数改进方案计算开销大,难以满足工业场景中对低延迟、低成本的实际需求,因此仍需进行轻量化<sup>[17-19]</sup>。

因此,本文在 ConvNeXt 架构基础上提出一种改进型图像识别模型,命名为 ConvNeXt\_CA。该模型以 ConvNeXt 在 ImageNet-1K 上获得的预训练权重作为初始化参数,结合本任务特性设计了分阶段参数冻结策略与注意力增强机制,在提升模型泛化能力的同时,增强其对目标区域的建模表达力。具体而言,在 ConvNeXt 的 Stage2 至 Stage5 的每个 Block 中,引入 CA 进行结构重构,并设计了轻量化融合结构 CANeXt Block,其核心思想是基于对蘑菇图像空间结构的理解,对 CA 模块进行定制性改造。考虑到蘑菇结构通常具有明显的几何对称性(如菌盖、菌褶的径向分布),本文将 CA 模块作为特征增强单元与主干网络卷积路径串联,引入残差式融合策略,其数学表达式为:

$$\mathbf{Y} = \text{DWConv}(\mathbf{X}) + \gamma \cdot \text{CA}(\mathbf{X}), \gamma \in \mathbf{R} \quad (1)$$

式中: $\mathbf{X}$  为输入特征,DWConv 表示 Depthwise 卷积操作,CA 为坐标注意力映射函数, $\gamma$  为可训练的融合权重系数。

该设计一方面保留了原始卷积特征的信息表达能力,另一方面允许模型根据训练过程自动调整注意力机制的贡献程度,在不显著增加计算开销的前提下,引入对目标区域的长程依赖建模能力。

为提高模型迁移效率并降低过拟合风险,本文采用“参数冻结、模块解冻”式训练流程。ConvNeXt\_CA 整体结构如图 1 所示。

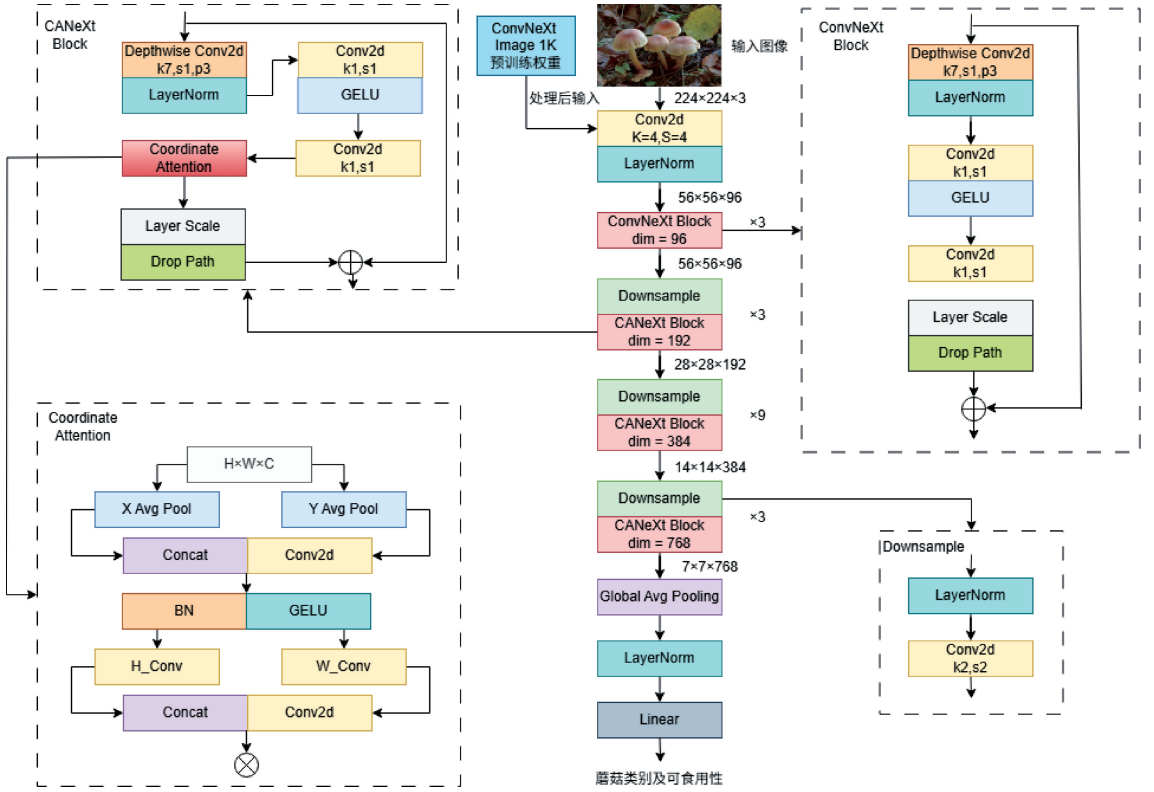


图 1 ConvNeXt\_CA 整体结构图

Fig. 1 Overall structure of ConvNeXt\_CA

## 1.2 迁移学习

相较于大规模数据集,蘑菇数据集规模较小,特征表示能力不足,因此,本文使用迁移学习技术来解决这一问题。迁移学习通过利用在大规模数据集上预训练的模型,能够借鉴已有特征表示,减少对庞大数据的依赖,同时提高训练效率和模型泛化能力。以 ConvNeXt 预训练模型为基础优化蘑菇识别。具体方法如下:

1) 采用 ConvNeXt 在 ImageNet-1K 上预训练,并获取相应权重;

2) 加载预训练权重,去除原有全连接层,设计适用于蘑菇识别的新全连接层,同时冻结除全连接层及 CA 模块外的其他网络层,以实现针对蘑菇识别任务的有效调整。

## 1.3 坐标注意力机制

在蘑菇识别任务中,ConvNeXt 结合了卷积神经网络(convolutional neural network, CNN)和 Transformer 的设计,保留了卷积的局部感知能力并增强了全局特征建模。然而,卷积操作缺乏对空间位置关系的精细建模,这对于形态差异微小且空间依赖性较强的蘑菇识别任务存在局限。为提升模型的空间感知能力,本文引入了坐标注意力机制 CA。CA 分为两个阶段:首先,沿水平方向与垂直方向对输入特征图  $\mathbf{X} \in \mathbf{R}^{C \times H \times W}$  进行全局平均池化,分别获得方向感知的特征表示,如式(2)、(3)所示。

$$\mathbf{Z}_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (2)$$

$$\mathbf{Z}_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (3)$$

接着,将  $\mathbf{Z}^h$  与  $\mathbf{Z}^w$  进行串联,并通过共享的  $1 \times 1$  卷积层,得到融合空间信息的中间特征表示,如式(4)所示。

$$\mathbf{f} = \delta(F_1([\mathbf{Z}^h, \mathbf{Z}^w])) \quad (4)$$

式中:  $\delta$  表示非线性激活函数,  $F_1$  为共享卷积操作。

随后,将  $\mathbf{f}$  沿空间维度分解为两个方向特征  $\mathbf{f}^h \in \mathbf{R}^{C/r \times H}$ 、 $\mathbf{f}^w \in \mathbf{R}^{C/r \times W}$ ,再分别通过两个  $1 \times 1$  卷积层映射回原始通道维度,得到注意力权重,如式(5)、(6)所示。

$$\mathbf{g}^h = \sigma(F_h(\mathbf{f}^h)) \quad (5)$$

$$\mathbf{g}^w = \sigma(F_w(\mathbf{f}^w)) \quad (6)$$

最终,通过逐点乘法将注意力权重施加于原始特征图,实现空间增强,如式(7)所示。

$$\mathbf{y}_c(i, j) = x_c(i, j) \times \mathbf{g}_c^h(i) \times \mathbf{g}_c^w(j) \quad (7)$$

## 1.4 知识蒸馏

尽管改进后的 ConvNeXt\_CA 在识别精度方面表现优异,但其庞大的模型参数和高计算开销限制了其在资源受限环境中的应用。为提升模型的部署效率,本文引入知识蒸馏方法,以 ConvNeXt\_CA 作为教师模型,将其知识迁移至轻量级的 ShuffleNet-V2 中,构建出新型网络 MushNet。该网络在保持较高识别精度的同时显著降低了计算成本与推理延迟,更适用于实时应用场景。

知识蒸馏核心思想在于通过教师模型的“软标签”引

导学生模型学习类间关系,而非仅依赖于真实标签。具体而言,训练好的教师模型  $T$  接收输入样本  $x$ , 输出各类别的 logits 向量  $z_T$ , 并经由温度系数  $\tau$  调整的 softmax 函数转化为概率分布,如式(8)所示。

$$P_{T,i} = \frac{\exp((Z_{T,i})/T)}{\sum_j \exp((Z_{T,j})/T)} \quad (8)$$

学生模型  $S$  同样输出 logits  $Z_s$ , 经 softmax 后得到对应的预测分布,如式(9)所示。

$$P_{S,i} = \frac{\exp((Z_{S,i})/T)}{\sum_j \exp((Z_{S,j})/T)} \quad (9)$$

学生模型  $S$  的训练目标是最大化其输出分布与  $T$  输出之间的差异,通常采用 Kullback-Leibler 散度(KL 散度)

衡量,如式(10)所示。

$$L_{KL} = \sum_{i=1}^n P_{T,i} \log(P_{T,i}/P_{S,i}) \quad (10)$$

此外,为保证学生模型  $S$  在真实标签上的分类能力,还需引入交叉熵损失,如式(11)所示。

$$L_{CE} = - \sum_{i=1}^n P_{true,i} \log(P_{S,i}) \quad (11)$$

最终总损失函数由两部分加权构成,如式(12)所示。

$$L_{total} = \alpha L_{KL} + (1 - \alpha) \cdot L_{CE} \quad (12)$$

通过最小化上述损失,学生模型 MushNet 不仅学习到教师模型的预测行为,还保留了对真实标签的判别能力,实现精度与效率的兼顾。知识蒸馏结构如图 2 所示。

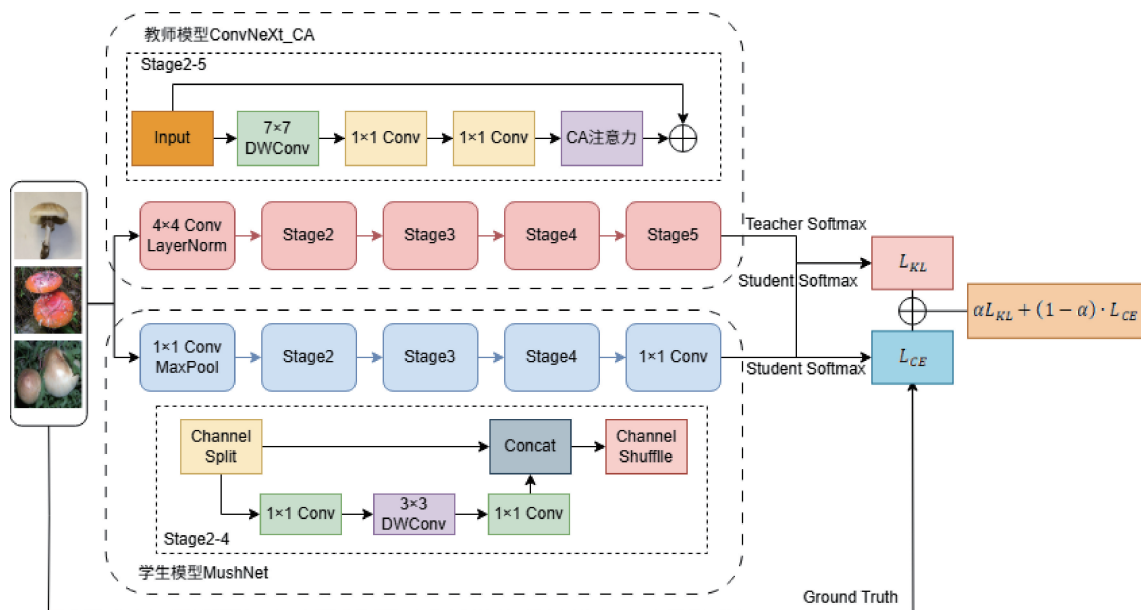


图 2 知识蒸馏结构图

Fig. 2 Knowledge distillation structure diagram

## 2 实验及分析

### 2.1 数据集与实验设置

#### 1) 数据集

本研究基于公开数据集 MO106。MO106 数据集涵盖 106 个类别,共 29 100 张蘑菇图像,各类别样本数量介于 105~581 张之间,平均约 275 张,且图像分辨率存在较大差异(最小 97 pixel×130 pixel,最大 640 pixel×640 pixel)。然而,该数据集存在两方面不足:一是蘑菇种类主要源自国外,缺乏对中国本土物种的覆盖;二是未对可食用蘑菇和有毒蘑菇进行明确区分。针对上述问题,本研究依据《中国毒蘑菇新修订名录》<sup>[20]</sup>和《中国菌根食用菌名录》<sup>[21]</sup>的权威分类标准,对 MO106 数据集进行了扩充和重构,构建了专门针对中国本土蘑菇的 MOCN 数据集。该数据集包含 52 个类别(其中可食用蘑菇 29 种,有毒蘑菇 23 种),

共计 14 630 张图像。这一优化不仅增强了数据集的地域适用性,同时提升了其在食用安全识别方面的实际应用价值。为提升数据集的均衡性和多样性,本研究采用了一系列图像增强策略,包括水平翻转、垂直翻转、角度旋转、亮度调整、噪声添加以及模糊处理等。针对样本量较少的类别,利用上述增强方法扩充数据,使每个类别的样本均达到约 500 张,最终构建出数据集 MOCN\_500 数据集共 26 182 张图像。数据增强后的数据集在类别均衡性、数据代表性以及泛化能力方面得到了显著优化。为了验证模型的泛化性,在此基础上还构建出 14 593 张图像的食用菌数据集 MOCN\_Edible 和 11 589 张图像毒蘑菇数据集 MOCN\_Poison。数据集部分样本量较少的类别图像增强前后对比如表 1 所示。

#### 2) 实验环境及参数设置

本实验在 Ubuntu 18.04 操作系统、NVIDIA GeForce



表 1 图像增强前后对比表

Table 1 Comparison of image before and after enhancement

张

序列号	类别	原始数量	增强后数量	序列号	类别	原始数量	增强后数量
1	暗蓝裸盖菇	264	504	13	赭红拟口蘑	268	500
2	丛生垂暮菇	278	502	14	虎掌刺银耳	272	504
3	大青褶伞	236	500	15	鳞皮扇菇	271	503
4	蝶形斑褶菇	220	500	16	绿盖鹅膏	321	505
5	灰光柄菇	214	502	17	毛柄网褶菌	300	500
6	美洲白环蘑	264	504	18	青绿湿伞	260	500
7	古巴裸盖菇	250	506	19	珊瑚状猴头	290	506
8	褐环乳牛肝菌	270	502	20	纹缘盔孢菌	234	506
9	环带斑褶菇	139	507	21	疣孢花边伞	232	504
10	皱环球盖菇	197	501	22	粘盖包脚菇	231	503
11	黄盖小脆柄菇	242	506	23	砖红垂幕菇	241	505
12	黄褐疣孢斑褶菇	153	505	24	蓝绿乳菇	187	507

RTX 3090 显卡、Python 3. 8、PyTorch 1. 11 和 CUDA 11. 4 环境下进行,旨在通过高效的硬件配置和深度学习框架,探索不同神经网络模型的性能表现。在实验过程中,采用了深度学习框架 PyTorch,构建并训练了多个神经网络模型,涵盖了从数据预处理到模型优化的多个环节。为了确保实验结果的公正性和变量的独立性,所有网络模型均采用了统一的超参数设置,这包括:输入图像尺寸为  $224 \times 224$ ,初始学习率设定为 0. 000 1,并通过余弦退火策略动态调整学习率,以优化训练过程中的收敛速度和稳定性。所有模型均使用了 AdamW 优化器,训练批次大小(batch size)设定为 32,损失函数选择交叉熵损失函数,以适应多分类任务的需求。训练过程共进行 100 个 Epoch,以确保模型能够充分学习数据中的特征。数据集的训练集与测试集按 8:2 的比例进行划分,确保了模型评估的可靠性和准确性。

3) 评价指标

为了评估网络性能,本文使用了 4 个指标:准确率(Accuracy)、精确率(Precision)、参数量(Params)和模型存储空间(Params Size)。

Accuracy 是衡量分类模型性能的基本指标,表示模型预测正确的样本数占总样本数的比例。其主要由正确预测为正类的样本数(TP)、正确预测为负类的样本数(TN)、错误预测为正类的样本数(FP)以及错误预测为负类的样本数(FN)4 个指标计算。其计算公式如式(13)所示。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

由于假阳性结果可能导致对有毒蘑菇的误判,从而引发潜在的健康或安全风险,因此在模型性能评估过程中,必须将精确率 Precision 作为核心评价指标之一,以确保正类预测结果的可靠性与实际应用的安全性。Precision 计算公式如式(14)所示。

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

在实际应用中,尤其是在资源受限的环境中,降低模型参数量和参数存储空间是一项重要的优化方向。因此,在模型设计与评估中,除了准确率等性能指标,还需综合考虑模型的参数量和存储空间,以实现更优的准确性-复杂度权衡。

Params 表示深度学习模型中的总参数量,通常用于衡量模型的规模和复杂度。Params Size 表示模型的可训练参数占用的存储空间,用于评估模型在存储受限设备上的可部署性。计算公式如式(15)、(16)所示。

$$Params = (K_w \times K_h \times C_{in} + bias) \times C_{out} \tag{15}$$

$$ParamsSize = \frac{Params \times Precision}{8 \times 10^6} \tag{16}$$

为了进一步提升模型的性能,同时兼顾模型的轻量化需求,本文采用了教师-学生模型架构的策略。该架构在保证高精度的前提下,通过将性能优秀的教师模型的知识迁移至结构更轻量的学生模型,实现了模型压缩与推理效率的提升。为此,本文首先对多种主流图像分类模型进行了对比实验,从中选取最优模型作为教师模型,配合知识蒸馏技术进行训练和优化,从而得到轻量化学生模型 MushNet。以下将分别介绍教师模型的选择、迁移学习有效性对比实验、注意力机制对比实验、消融实验,以及学生模型 MushNet 的设计与性能评估。

2. 2 教师模型实验

1) 教师模型对比实验

为验证所提教师模型在蘑菇图像识别中的泛化能力与有效性,本文选取了近年来具有代表性的 4 种主流图像分类网络作为对比模型,包括 Swin-Transformer<sup>[22]</sup>、ResNeSt<sup>[23]</sup>、FocalNet<sup>[24]</sup> 和 ConvNeXt。这些模型涵盖了注意力机制、多尺度融合、动态卷积等主流架构思想,能够全面体现所提方法的性能优势。实验在 MOCN\_Edible、

MOCN\_Poison 和 MOCN\_500 三个蘑菇数据集上进行,结果如表 2~4 所示。

表 2 MOCN\_Edible 对比结果  
Table 2 MOCN\_Edible comparison results

模型名称	Accuracy/%	Precision/%	Params/M	ParamsSize/MB
Swin-Transformer	85.43	85.68	27.56	114.26
ResNeSt	86.82	87.01	25.54	110.04
FocalNet	83.88	84.26	28.65	113.64
ConvNeXt	86.77	87.09	27.85	106.25
ConvNeXt_CA	95.98	96.02	28.18	107.49

表 3 MOCN\_Poison 对比结果  
Table 3 MOCN\_Poison comparison results

模型名称	Accuracy/%	Precision/%	Params/M	ParamsSize/MB
Swin-Transformer	86.29	86.58	27.56	114.26
ResNeSt	87.54	87.59	25.54	110.04
FocalNet	85.46	85.51	28.65	113.64
ConvNeXt	87.83	87.95	27.85	106.25
ConvNeXt_CA	96.45	96.50	28.18	107.49

表 4 MOCN\_500 对比结果  
Table 4 MOCN\_500 comparison results

模型名称	Accuracy/%	Precision/%	Params/M	ParamsSize/MB
Swin-Transformer	85.63	85.81	27.56	114.26
ResNeSt	87.15	87.31	25.54	110.04
FocalNet	83.72	83.86	28.65	113.64
ConvNeXt	87.43	87.65	27.85	106.25
ConvNeXt_CA	94.89	94.96	28.18	107.49

表 2、3 和表 4 的实验结果表明,本文提出的 ConvNeXt\_CA 模型在 3 个数据集上的分类精度均优于其他对比模型,体现了本文方法在性能上的显著提升,也进一步验证了其良好的泛化能力。

考虑到实际应用中蘑菇识别需同时判断种类与毒性,模型目标应为全类别准确分类,而非仅在已知毒/食用前提下的细粒度识别。因此,后续实验选用类别覆盖更广、扩展性更强的 MOCN\_500 数据集作为主要数据集,以全面验证所提方法的有效性 with 适用性。

2) 迁移学习有效性对比实验

为了验证迁移学习方法对模型性能以及收敛有促进作用,以 ConvNeXt\_CA 模型为参照,对比 3 种训练方式下验证集的准确率以及收敛过程,3 种训练方式的实验分别如下:

实验 1<sup>\*\*\*</sup>: 不加载 ImageNet 预训练权重进行训练。

实验 2<sup>\*\*\*</sup>: 解冻所有网络层,设置所有特征提取层参数可训练。

实验 3<sup>\*\*\*</sup>: 冻结除 CA 和全连接层外的所有网络层,仅

训练 CA 和全连接层参数。

3 种训练方式下模型在验证集的准确率结果如表 5 所示。

表 5 迁移学习有效性对比  
Table 5 Comparison of the effectiveness of transfer learning

实验编号	训练方式	Accuracy/%
1 <sup>***</sup>	不加载预训练权重	88.60
2 <sup>***</sup>	训练所有网络层	91.86
3 <sup>***</sup>	冻结/解冻网络层	94.89

3 种训练方式下模型在验证集的准确率和收敛情况对比如图 3 所示。

如表 5 和图 3 所示,不加载预训练权重时,模型收敛慢、准确率波动大且最终精度较低;若解冻全部网络层进行训练,收敛速度有所提升,但仍存在一定波动;而本文采用的迁移学习策略表现出更快的收敛速度和更稳定的训练过程,最终取得了最高准确率,表明该方法能有效提升模型性能并加快收敛。

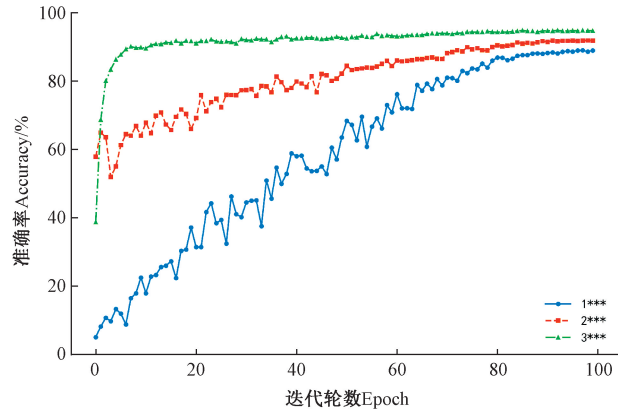


图 3 不同训练方式下模型准确率对比

Fig. 3 Comparison of model accuracy under different training methods

3) 注意力对比实验

为了验证 CA 机制的优势,在加载预训练权重的 ConvNeXt 模型中,分别引入高效通道注意力<sup>[25]</sup>(efficient channel attention, ECA)、卷积块注意力机制<sup>[26]</sup>(convolutional block attention module, CBAM)、高效多尺度注意力<sup>[27]</sup>(efficient multi-scale attention, EMA)和 SimAM<sup>[28]</sup>四种注意力机制进行对比实验,实验结果如表 6 所示。

实验结果表明,所有注意力机制均提升了模型准确率,其中 CA 提升最为显著,达到 94.89%,且仅小幅增加了模型参数,综合表现在所有注意力中最优。

4) 消融实验

为了验证本文提出方法对 ConvNeXt 模型性能的提升效果,本文通过消融实验对各部分优化方法进行了测试,结果如表 7 所示。

表 6 注意力模块对比实验

Table 6 Comparative experiment of attention module

注意力模块	Accuracy/ %	Params/ M	ParamsSize/ MB
ECA	93.39	37.56	143.27
CBAM	92.34	31.09	118.60
EMA	92.72	27.88	106.38
SimAM	88.54	27.85	106.25
CA	94.89	28.18	107.49

表 7 消融实验结果

Table 7 Ablation experiment results

方法	Accuracy/ %	Params/ M	ParamsSize/ MB
ConvNeXt	87.43	27.85	106.25
ConvNeXt+迁移学习	88.89	27.85	106.25
ConvNeXt+CA	88.60	28.18	107.49
Ours	<b>94.89</b>	<b>28.18</b>	<b>107.49</b>

实验结果表明,在原始模型的基础上,迁移学习和引入 CA 注意力机制均使模型的准确率略微提升,而本文所提出的方法在此基础上显著提高了准确率,提升幅度达 7%,且参数数量和存储空间仅有轻微上升,这进一步证明了本文方法在蘑菇图像识别任务中能够显著提升模型性能。

2.3 学生模型对比实验

为了验证 MushNet 在轻量化模型中的精度优势,本文针对 MushNet 做了与其他广泛应用于图像分类的轻量化模型的对比实验,包括 MobileNet-V3<sup>[29]</sup>、RegNet<sup>[30]</sup>、EfficientNet-V2<sup>[31]</sup>以及 ShuffleNet-V2,结果如表 8 所示。

表 8 学生模型对比结果

Table 8 Comparison results of student models

模型名称	Accuracy/%	Precision/%	Params/M	ParamsSize/MB
MobileNet-V3	81.80	81.93	1.57	5.99
RegNet	82.28	82.48	2.33	8.91
EfficientNet-V2	86.72	86.84	20.24	77.23
ShuffleNet-V2	88.64	88.78	1.31	4.99
MushNet(Ours)	<b>91.42</b>	<b>91.51</b>	<b>1.31</b>	<b>4.99</b>

实验结果表明,本文提出的 MushNet 在精度上均优于其他轻量化模型,证明知识蒸馏能够在维持模型轻量化的基础上,有效提升其精度表现。

2.4 知识蒸馏实验

在知识蒸馏过程中,温度系数  $T$  和权重  $\alpha$  是两个非常关键的超参数。温度系数主要用于平滑教师模型的输出概率分布,使学生模型能够更好地从教师模型中学习丰富的知识。为了探索最佳温度系数,本文针对温度系数  $T$

进行对比实验,结果如表 9 所示。

由表 9 可知,温度系数  $T$  在取值为 3 时,准确率最高(90.42%)。温度系数继续增大时,准确率开始有所下降,这表明进一步增大温度系数  $T$  可能会使得输出分布过于平滑,导致学生模型对某些类别的区分度降低。因此温度系数  $T$  最终取值为 3。

训练过程中,学生模型既要从小教师模型的预测中学习(软目标),又要保持与真实标签的准确性(硬目标)。通过

表 9 温度系数对比实验

Table 9 Temperature coefficient comparison experiment

温度系数 $T$	Accuracy/%
3	90.42
4	90.34
5	90.10
6	89.92
7	90.02
8	90.23

调整  $\alpha$ , 可以控制这两部分损失的相对重要性。为了确定最佳权重参数  $\alpha$ , 在温度系数  $T$  取值为 3 的基础上, 进行权重参数  $\alpha$  对比实验, 从而确保学生模型发挥最大性能。结果如表 10 所示。

由表 10 可知, 权重参数  $\alpha$  取值为 0.8 时学生模型达到最高准确率。

综上, 当温度系数  $T$  设置为 3、权重参数  $\alpha$  设定为 0.8 时, 学生模型在知识蒸馏过程中取得了最优性能表现。

2.5 与已有蘑菇识别方法对比

近年来, 基于深度学习的蘑菇识别方法在提升准确率

表 10 权重参数对比实验

Table 10 Comparison experiment of weight parameter

权重参数 $\alpha$	Accuracy/%
0.4	89.58
0.5	90.42
0.6	90.25
0.7	90.79
0.8	91.42
0.9	91.30

的同时, 逐步向轻量化方向发展, 以适应移动端部署需求。王江晴等<sup>[32]</sup>提出 BE-EfficientNet, 融合双注意力机制, 在提升性能的同时降低模型复杂度; 钱嘉鑫等<sup>[33]</sup>改进 CBAM 模块, 设计 PA\_CBAM 并优化 ResNet50 结构, 同时构建了轻量级 MobileNet\_v2+PA\_CBAM 模型用于移动端应用; 周罕觅等<sup>[34]</sup>虽未针对蘑菇识别, 但其提出的 LWCR-Net 在植物病害任务中采用知识蒸馏与轻量网络结合的设计思路, 与本文方法具有一定相似性, 因此也被纳入对比。

本文所提出的方法与上述几种模型在数据集 MOCN\_500 上进行了对比实验, 结果如表 11 所示。

表 11 与已有方法对比

Table 11 Comparison with existing methods

模型名称	Accuracy/%	Precision/%	Params/M	ParamsSize/MB
BE-EfficientNet	88.93	89.02	4.55	17.35
PA-CBAM-ResNet	87.11	87.25	28.64	109.28
PA-CBAM-pretrain-MobileNetV2	82.13	82.20	2.32	8.84
LWCR-Net	88.93	89.04	1.65	6.30
Ours	<b>91.42</b>	<b>91.51</b>	<b>1.31</b>	<b>4.99</b>

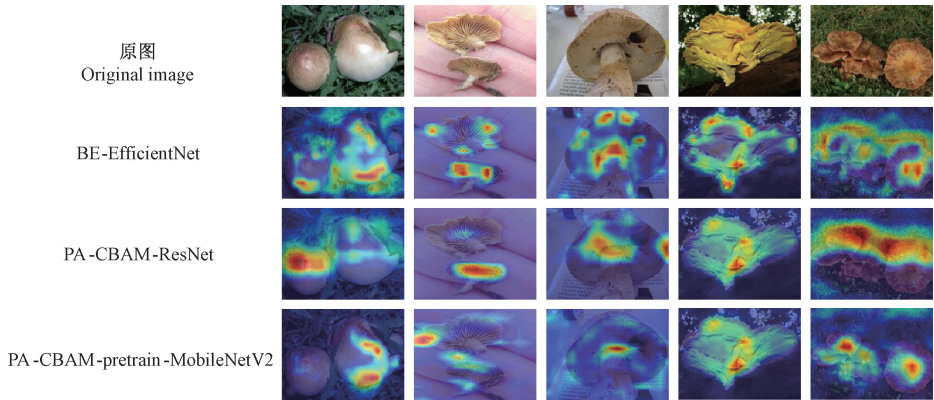
由表 11 可以看出, 本文在准确率(Accuracy)和精确率(Precision)指标上均优于现有方法, 且参数数量和模型大小均为最低, 展现出更优的识别性能与轻量化优势, 适用于资源受限的边缘设备部署场景。

2.6 蘑菇图像可视化

为了提升模型的可解释性, 本文采用了基于梯度的类激活映射 Grad-CAM 方法, 对模型在推理过程中所提取的

关键特征进行可视化处理。本文选取古巴裸盖菇、鳞皮扇菇、大紫蘑菇、硫磺菌和美洲白环蘑 5 类具有代表性的蘑菇样本进行实验, 与已有蘑菇识别模型和相似任务模型的热力图进行对比, 结果如图 4 所示。

结果表明, MushNet 在热力图中展现出更清晰且集中于目标区域的响应特征, 关注区域更符合人类视觉识别的直觉判断, 进一步验证了其在轻量化模型中的优越性能。





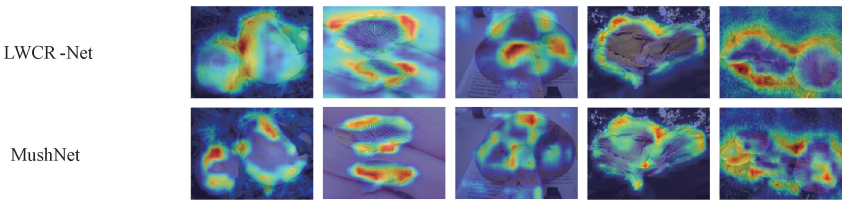


图 4 各模型特征热力图对比

Fig. 4 Comparison of heatmaps of various model features

3 结 论

本研究针对自然环境下蘑菇图像识别过程中存在的类别众多、形态相似、背景复杂以及模型部署受限等关键问题,提出了一种结合改进型 ConvNeXt 结构与知识蒸馏技术的识别方法。通过迁移学习引入 ImageNet 预训练权重,加速了模型在特定任务上的收敛过程并提升了其泛化能力;在基础网络中引入坐标注意力机制(CA),有效强化了对细粒度特征的表征能力,显著提升了模型对复杂背景下蘑菇图像的识别准确率;同时,设计了以 ConvNeXt\_CA 作为教师模型、ShuffleNet V2 作为学生模型的蒸馏框架,在保持较高识别性能的基础上,大幅压缩了模型体积,降低了推理开销,最终构建出适合移动端部署的轻量化蘑菇识别模型 MushNet。综合实验结果表明,该方法在准确率与模型效率之间实现了良好平衡,具备在实际应用中推广的可行性与实用价值,为毒蘑菇误食风险的早期预警与食品安全保障提供切实可行的技术支撑。

参考文献

[1] 江仿. 野生食用菌资源的可持续开发策略[J]. 中国食用菌,2020,39(6):89-92.  
JIANG F. Sustainable development strategy of wild edible fungi resources[J]. Edible Fungi of China, 2020, 39(6): 89-92.

[2] 赵碧钗. 误食毒蘑菇后中毒的护理[J]. 智慧健康, 2024,10(25):182-184,188.  
ZHAO B CH. Nursing for poisoning after ingestion of poisonous mushrooms [J]. Smart Health, 2024, 10(25): 182-184, 188.

[3] 高帆,谭廷鸿,孙琰妮,等. 毒蘑菇鉴别技术研究进展[J]. 食品安全质量检测学报,2024,15(9):51-61.  
GAO F, TAN T H, SUN Y N, et al. Research progress in identification technology of poisonous mushrooms[J]. Journal of Food Safety and Quality, 2024, 15(9): 51-61.

[4] PENG Y, XU Y, SHI J, et al. Wild mushroom classification based on improved mobilevit deep learning[J]. Applied Sciences, 2023, 13(8): 4680.

[5] CENGİL E, ÇINAR A. Poisonous mushroom detection using YOLOV5 [J]. Turkish Journal of

Science and Technology, 2021, 16(1): 119-127.

[6] 王莉,董鹏豪,王瞧,等. 基于改进 ResNet18 的干香菇等级识别 [J]. 国外电子测量技术, 2024, 43 (1): 117-125.  
WANG L, DONG P H, WANG Q, et al. Dried shiitake mushroom grade recognition based on improved ResNet18 [ J ]. Foreign Electronic Measurement Technology, 2024, 43(1): 117-125.

[7] 姚芷馨,张太红,赵昀杰. 基于改进权重衰减的 EfficientNet 食用菌图像识别[J]. 食品与机械,2022, 38(11):117-124.  
YAO ZH X, ZHANG T H, ZHAO Y J. Edible fungus image recognition based on EfficientNet with improved weight decay [J]. Food and Machinery, 2022, 38(11): 117-124.

[8] 张志刚,余鹏飞,李海燕,等. 基于多尺度特征引导的细粒度野生菌图像识别[J]. 激光与光电子学进展, 2022,59(12):182-191.  
ZHANG ZH G, YU P F, LI H Y, et al. Fine-grained wild mushroom image recognition based on multi-scale feature guidance [ J ]. Laser & Optoelectronics Progress, 2022, 59(12): 182-191.

[9] LIU Z, MAO H Z, WU C Y, et al. A convnet for the 2020s [ C ]. IEEE/CVF Conference on Computer Vision and Pattern Recognition,2022:11976-11986.

[10] HOU Q B, ZHOU D Q, JIA SH F. Coordinate attention for efficient mobile network design[C]. CVF Conference on Computer Vision and Pattern Recognition,2021:13708-13717.

[11] 丁庆松,孙昊,李强,等. 基于迁移学习的肠衣质量检测[J]. 电子测量技术,2023,46(11):185-192.  
DING Q S, SUN H, LI Q, et al. Casing quality detection based on transfer learning [J]. Electronic Measurement Technology, 2023,46(11): 185-192.

[12] 王欣,邓章俊,秦斌. 基于轻量化网络 and 知识蒸馏的回转窑工况识别[J]. 电子测量与仪器学报,2023,37(9): 149-159.  
WANG X, DENG ZH J, QIN B. Rotary kiln working condition identification based on lightweight network and knowledge distillation [ J ]. Journal of Electronic

- Measurement and Instrumentation, 2023, 37(9): 149-159.
- [13] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: Practical guidelines for efficient CNN architecture design [C]. Computer Vision-ECCV 2018: 15th European Conference, 2018: 122-138.
- [14] 杨宇轩, 钟宝荣. 基于改进 ConvNeXt 模型的轻量化页岩岩性识别方法[J]. 现代电子技术, 2024, 47(24): 131-136.
- YANG Y X, ZHONG B R. Lightweight lithologic identification method of cuttings based on improved ConvNeXt model[J]. Modern Electronic Technology, 2024, 47(24): 131-136.
- [15] 杨德龙, 李婧. 基于注意力与小平方核的 ConvNeXt 农业杂草识别方法[J]. 江苏农业科学, 2024, 52(14): 207-214.
- YANG D L, LI J. ConvNeXt agricultural weed identification method based on attention and small square kernel [J]. Jiangsu Agricultural Sciences, 2024, 52(14): 207-214.
- [16] 田甜, 程志友, 鞠薇, 等. 基于 SimAM-ConvNeXt-FL 的茶叶病害小样本分类方法研究[J]. 农业机械学报, 2024, 55(3): 275-281.
- TIAN T, CHENG ZH Y, JU W, et al. Research on small sample classification of tea diseases based on SimAM-ConvNeXt-FL [J]. Journal of Agricultural Machinery, 2024, 55(3): 275-281.
- [17] 彭军强, 张立坤, 杨亚楠. 基于多模态轻量化混合模型的情绪识别[J]. 电子测量技术, 2024, 47(3): 9-18.
- PENG J Q, ZHANG L K, YANG Y N. Emotion recognition based on multimodal lightweight hybrid model [J]. Electronic Measurement Technology, 2024, 47(3): 9-18.
- [18] 张博伟, 何彦霖, 王康, 等. 基于全方位深层加权轻量化网络的冠脉造影图像超分辨率重建方法[J]. 仪器仪表学报, 2024, 45(7): 200-209.
- ZHANG B W, HE Y L, WANG K, et al. Super resolution reconstruction method of coronary angiography image based on omnidirectional deep weighted lightweight network[J]. Chinese Journal of Scientific Instrument, 2024, 45(7): 200-209.
- [19] LI H, HE L L, HE N, et al. A two-stage sheep facial pain recognition method based on deep learning[J]. Instrumentation, 2023, 10(3): 42-52.
- [20] 图力古尔, 李海蛟, 包海鹰, 等. 中国毒蘑菇新修订名录[J]. 菌物研究, 2024, 22(4): 301-321.
- BAU T, LI H J, BAO H Y, et al. New revised list of poisonous mushrooms in China [J]. Mycological Research, 2024, 22(4): 301-321.
- [21] 魏杰, 高巍, 黄晨阳. 中国菌根食用菌名录[J]. 菌物学报, 2021, 40(8): 1938-1957.
- WEI J, GAO W, HUANG CH Y. List of Chinese mycorrhizal edible fungi [J]. Mycological Journal, 2021, 40(8): 1938-1957.
- [22] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]. IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 10012-10022.
- [23] ZHANG H, WU CH R, ZHANG ZH Y, et al. Resnest: Split-attention networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 2736-2746.
- [24] YANG J W, LI C Y, DAI X Y, et al. Focal modulation networks [C]. Advances in Neural Information Processing Systems, 2022: 4203-4217.
- [25] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. CVF Conference on Computer Vision and Pattern, 2020: 11531-11539.
- [26] SANGHYUN W, JONGCHAN P, JOON Y L, et al. CBAM: Convolutional block attention module [C]. European Conference on Computer Vision (ECCV), 2018: 3-19.
- [27] OUYANG D L, HE S, ZHANG G ZH, et al. Efficient multi-scale attention module with cross-spatial learning [C]. 2023 IEEE International Conference on Acoustics, Speech and Signal Processing: ICASSP 2023, 2023: 776-780.
- [28] YANG L X, ZHANG R Y, LI L, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks [C]. International Conference on Machine Learning: ICML 2021, 2022: 11853-11864.
- [29] ANDREW H, MARK S, GRACE C, et al. Searching for MobileNetV3 [C]. 2019 IEEE/CVF International Conference on Computer Vision: ICCV 2019, 2019: 1314-1324.
- [30] ILIJA R, RAJ P K, ROSS G, et al. Designing network design spaces [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition: CVPR 2020, 2020: 10425-10433.
- [31] TAN M X, LE Q. EfficientNetV2: Smaller models and faster training [C]. International Conference on Machine Learning: ICML 2021, 2022: 10086-10096.
- [32] 王江晴, 马春, 莫海芳, 等. 基于融合双注意力机制的野生菌图像识别方法[J]. 中国农机化学报, 2025, 46(6): 70-76.

WANG J Q, MA CH, MO H F, et al. Wild mushroom image recognition method based on fusion dual attention mechanism [J]. China Agricultural Machinery Chemistry News, 2025, 46(6): 70-76.

[33] 钱嘉鑫,余鹏飞,李海燕,等. 基于特征融合与注意力机制的野生菌细粒度分类[J]. 激光与光电子学进展, 2023,60(4):110-119.

QIAN J X, YU P F, LI H Y, et al. Fine granularity classification of wild fungi based on feature fusion and attention mechanism [J]. Progress in Laser and Optoelectronics, 2023,60(4): 110-119.

[34] 周罕觅,陈佳庚,代智光,等. 基于知识蒸馏和轻量级卷积神经网络的植物病害识别方法[J]. 南京农业大学学报,2024,47(6):1189-1201.

ZHOU H M, CHEN J G, DAI ZH G, et al. Plant disease recognition method based on knowledge distillation and lightweight convolution neural network[J]. Journal

of Nanjing Agricultural University, 2024, 47(6): 1189-1201.

**作者简介**

任喜伟(通信作者),博士,教授级高级工程师,主要研究方向为机器视觉、过程系统工程等。  
E-mail:renxiwei@sust.edu.cn

王瑞,硕士研究生,主要研究方向为深度学习、图像处理等。  
E-mail:wangrui124@qq.com

贾事端,硕士研究生,主要研究方向为计算机视觉、目标检测等。  
E-mail: 2941149177@qq.com

刘艳,硕士研究生,主要研究方向为计算机视觉、视频摘要等。  
E-mail: 231612057@sust.edu.cn

肖曼,硕士研究生,主要研究方向为计算机视觉。  
E-mail: 1724648452@qq.com